Lawrence Berkeley National Laboratory

SARS-CoV-2/COVID-19 MULTIDISCIPLINARY RESEARCH STRATEGIC PLAN

DECEMBER 2020

BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

EXECUTIVE SUMMARY

In late 2019 and throughout 2020, the emergence of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), its resultant coronavirus disease 2019 (COVID-19), and the explosion of the global pandemic rapidly mobilized scientists and research institutions across the world to focus on developing a thorough understanding of the virus and the disease and to translating that knowledge into effective diagnostics, treatments and mitigation strategies. The United States government supported research at its 17 Department of Energy National Laboratories through the Coronavirus Aid, Relief and Economic Security (CARES) Act, dedicating nearly $100 million to efforts in the National Laboratory complex. As part of these efforts and others, Lawrence Berkeley National Laboratory quickly built a significant body of research addressing the challenges posed by SARS-CoV-2 and COVID-19. In the nine months since the pandemic was declared, many of the questions surrounding the virus and the disease have crystallized, enabling a more strategic approach to future research that can reduce the impacts and duration of the pandemic.

Berkeley Lab, in the spirit of its mission to bring science solutions to the world, developed a strategic plan for a team science approach for additional SARS-CoV-2 and COVID-19 research, focusing on the key roles for National Laboratories in the U.S. research ecosystem. In this SARS-CoV-2/COVID-19 Multidisciplinary Research Strategic Plan, strategies for new efforts are grouped into platform technologies and novel research. In both of these sections of the plan, Berkeley Lab's strength in multidisciplinary, large-scale research is a foundation that will enable advances in the understanding of SARS-CoV-2 transmission, prevent future infections, treat COVID-19, mitigate additional transmission, provide insights for possible policy interventions, and develop new technologies that will enable accelerated discovery in all these areas.

The ideas for new platform technologies described in this strategic plan focus on high-throughput methods development, integration of multiple technologies, and application of advanced computational methods to challenging problems. The four strategies for platform technologies described later in this plan are:

- Strategy for a biological systems characterization facility
- Strategy for integration of structural biology and imaging methods
- Strategy for high-throughput materials discovery and characterization, including new materials, those made by biological organisms, and bio-inspired materials
- Strategy for microsystems, robotics, automation, and self-driving labs.

The three new research strategies, all of which would involve large research teams that leverage Berkeley Lab's supercomputing capabilities, are:

- Strategy for inverse design of compounds
- Strategy for ecosystem surveillance
- Strategy for understanding the role of human mobility in transmission of SARS-CoV-2.

With this plan, Berkeley Lab has developed an approach for SARS-CoV-2 and COVID-19 research that is inclusive of the diversity of discovery and applied science across the lab and that readies us for future opportunities to solve critical national and global challenges beyond this current pandemic.

TABLE OF CONTENTS

INTRODUCTION

The emergence of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and resultant coronavirus disease 19 (COVID-19) in late 2019 mobilized scientists across the globe to undertake research needed to combat the virus and to treat the disease. Through the Coronavirus Aid, Relief and Economic Security (CARES) Act, the United States government marshaled its scientific research agencies to support research on SARS-CoV-2 and COVID-19. The U.S. Department of Energy (DOE) Office of Science received $99.5 million and established the National Virtual Biotechnology Laboratory (NVBL) to leverage its 17 National Laboratories and their extensive capabilities for SARS-CoV-2 and COVID-19 research. National Laboratories have also been resources and scientific partners for other institutions in the race to stop the spread of SARS-CoV-2 and mitigate the effects of COVID-19. As part of the NVBL, Berkeley Lab leveraged its key strengths in biological, chemical, computational, materials, and physics research, its User Facilities, and its core capabilities in mobility and buildings research, to develop solutions through a multidisciplinary team science approach.

While the pandemic spurred scientists to quickly respond to this emergency, we cannot sustain such activity without a set of priorities and a strong plan for implementation. It seems clear that COVID will be endemic, and Berkeley Lab has developed a vision for how we can address future challenges. To develop this plan, an inclusive survey was sent through a variety of distribution methods to capture the scientific challenges posed by the pandemic and to gather ideas from staff across the lab in all job categories. The survey results informed two half-day visioning sessions attended by around 40 participants. These visioning sessions were focused on 1) articulating visions for SARS-CoV-2 and COVID-19 research and 2) how the expertise, capabilities, and infrastructure developed through that research can be translated to enable DOE (and other sponsors') mission science.

This strategic plan represents the input of more than 100 Berkeley Lab staff, provides suggestions for projects that can be started quickly, and includes lessons learned and recommendations for teaming based on experience with NVBL and other research programs.

SCIENTIFIC CHALLENGES POSED BY THE SARS-CoV-2 VIRUS AND COVID-19 PANDEMIC

In a survey sent to Berkeley Lab staff, respondents were asked to identify scientific challenges posed by the pandemic that could be addressed through coordinated research at Berkeley Lab and across the network of National Laboratories, including public-private partnership with industry and other federal and state agencies. The challenges suggested by respondents are grouped into the three categories below.

Rapid, deployable testing

One of the major challenges resulting from the shock of a quickly growing pandemic is the lack of availability of rapid, sensitive, specific and quickly deployable testing methods to detect the presence of SARS-CoV-2 virus in humans and their surroundings. Disruption of global supply chains and increased demand have limited the availability of common testing reagents and consumables, while methods of collection require trained staff, designated testing areas, and equipment that minimizes spread of the virus. Research into new testing methods, reagents, consumables and protective equipment could alleviate some of the current challenges with testing. Many new methods that can reliably detect the virus in easily acquired samples, for example from saliva, are under investigation and efforts could be expanded to understand the minimum, simplest and most cost-effective sampling approach necessary. Methods for testing samples for presence of the virus could be scaled down into miniaturized systems that allow for rapid assays using minimal reagents, requiring new research in microsystems development. Sample pooling, for example, also offers a straightforward and economic approach to scale up testing, provided the efficacy of this method can be established. Tests that can simultaneously survey for both the virus and antibodies would allow for the assessment of an individual's status at a point in time. Research into the types of samples used for testing and new assays for dual detection could fulfill that aspiration. Through the discovery and development of new materials that have antiviral properties and could be easily cleaned or sterilized with minimum risk of exposing others to SARS-CoV-2, testing could be self-administered and performed outside of clinical settings.

Berkeley Lab researchers, along with those across the DOE National Laboratory complex, have deep expertise, capabilities, and facilities that can contribute to efforts to speed distributed testing for SARS-CoV-2 virus and antibodies. Leveraging the advanced genome sciences capabilities at the Joint Genome Institute at Berkeley Lab, scientists have developed new methods of pooling samples that allow for quicker screening of multiple samples and use less of necessary reagents and consumables. Additionally, researchers are working to develop microsystems, miniaturized testing and analytical platforms, that can replicate laboratory procedures in very small volumes, to further reduce the amounts of needed reagents and materials needed for testing for SARS-CoV-2.

Development of therapeutics and treatments

Researchers around the world in academia, industry, and at National Labs are racing to understand the mechanisms behind infection of SARS-CoV-2, how the virus interacts inside of host cells, and how those processes and interactions may be disrupted to design new therapeutics. Research into the structure of

SARS-CoV-2 proteins and those of the host that interact with SARS-CoV-2 proteins, as well as the mechanisms underlying those interactions, will enable researchers to develop targeted therapeutics that reduce the severity of COVID-19 or possibly prevent disease altogether.

At Berkeley Lab's Advanced Light Source, and other DOE light sources, researchers are studying viral proteins to understand their structure and mechanisms of interaction with human proteins to gain new insights into how these proteins could be targets for antiviral medications or as targets for therapeutics that treat the effects of COVID-19. New algorithms and methodologies for artificial intelligence (AI) and machine learning developed at Berkeley Lab are capable of gleaning insights from the scientific literature on SARS-CoV-2 and COVID-19, knowledge that can be applied for inverse design of currently unknown compounds that can be used as antivirals or treatments.

Understanding and mitigating the spread of SARS-CoV-2

One of the challenges with tackling a new human pathogen is understanding how it is spread, under what conditions, and how that spread can be mitigated. Further research is needed to understand the properties of viral particles and how they might be most effectively contained, perhaps with novel materials as barriers in face coverings, in air handling units, and on surfaces. Additionally, investigation into the modes of transmission, seasonal effects, and how humans interact within built structures or transportation infrastructure will be critical to resuming activities indoors. New materials research and development for materials that are anti-viral or that can be easily sterilized will be needed to minimize spread of SARS-CoV-2.

Researchers at Berkeley Lab have modeled population mobility and ventilation to address critical questions around energy usage, and now these models can be combined with epidemiological and viral spread models to understand how SARS-CoV-2 is transmitted in buildings and in various transportation scenarios. The expertise and capabilities for the discovery and development of new materials at the Molecular Foundry can be applied to develop new materials and coatings that indicate contamination with SARS-CoV-2 or that can be used to trap and neutralize viral particles in areas where there is high likelihood of spread, such as building ventilation systems.

ABOUT THE STRUCTURE OF THE STRATEGIC PLAN

During Berkeley Lab's internal visioning process that formed the basis for this strategic plan, participants discussed a wide variety of visionary and actionable approaches to addressing the scientific challenges posed by SARS-CoV-2 and COVID-19. Berkeley Lab prides itself on providing science solutions to the world through a team science approach. To align with this ideal, concepts from the visioning process were synthesized into seven strategies. These seven strategies were then grouped into two sections representing Berkeley Lab's core strengths in the development of platform technologies and multidisciplinary research. The core strengths and strategies represent the unique role that the National Laboratories occupy in the U.S. research ecosystem: as a place for large-scale instrumentation and the development of future research infrastructure, and as a home for research teams that cross the boundaries of scientific disciplines to deliver new breakthroughs and the seeds for future research to address national-scale challenges.

STRATEGIES FOR NEW PLATFORM TECHNOLOGIES TO ADVANCE SARS-CoV-2/ COVID-19 AND DOE MISSION RESEARCH

STRATEGY FOR A BIOLOGICAL SYSTEMS CHARACTERIZATION FACILITY

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

Experience with the NVBL response to the COVID-19 pandemic shows that developing therapeutics and new testing approaches relies on a detailed understanding of many aspects of the virus: genome, biochemical and cellular functions of key proteins, interactions with host proteins, and life cycle. Fortunately, the *Coronaviridae,* the family of which SARS-CoV-2 belongs, has been studied for many years around the world, and existing projects were able to pivot quickly to execute detailed characterizations of the newly emerged virus. While we had a bit of a head start with COVID-19, in the future, we will likely be dealing with completely new viruses, and the BioSCF will help to characterize new pathogens.

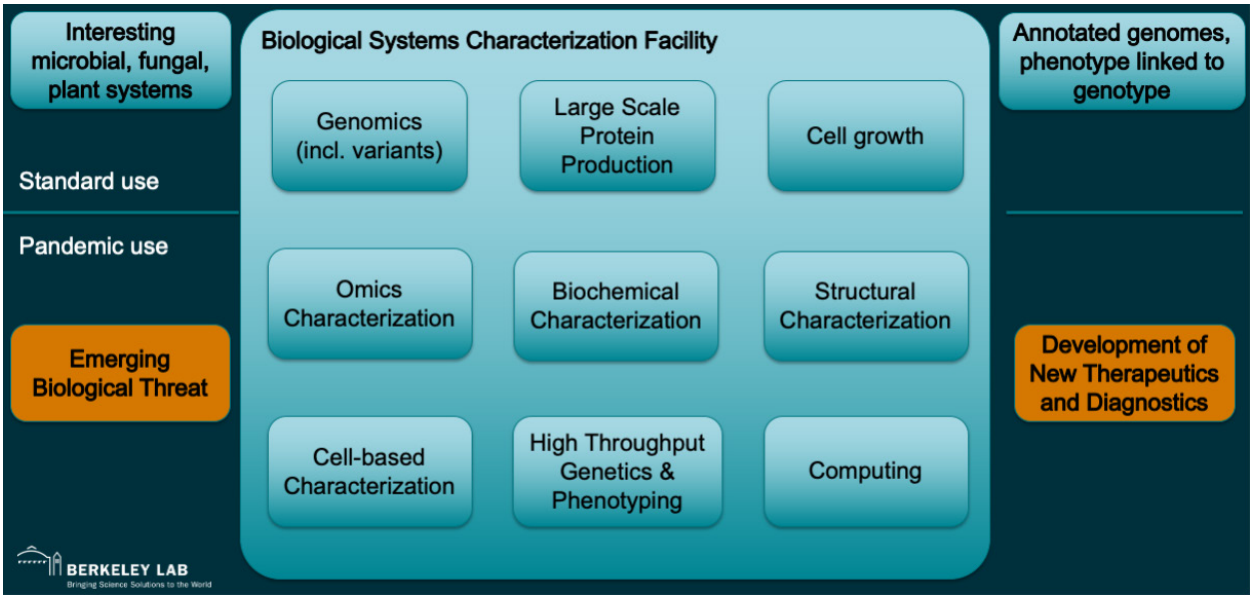


Figure 1: Scope of the Biological Systems Characterization Facility.

The development of a facility that is able to rapidly characterize emerging pathogens will be a key weapon in the fight against future biological threats that have the potential to become pandemics. The same facility will also be a critical research tool in advancing an understanding of microbes and viruses that are important to DOE's Office of Biological and Environmental Research (BER), particularly the Biological Systems Science Division, mission. The newly proposed Biological Systems Characterization Facility (BioSCF, **Figure 1**) will, for the first time, bring together LBNL expertise in genomics, protein production, omics, high throughput genomics and biochemistry, structural biology, imaging and computing. This will be a unique national resource for the rapid and detailed characterization of gene products from any organism, including viruses, microbes, algae, protists, and plants. It will be a key tool in

addressing DOE's Office Biological and Environmental Research (BER) grand challenge of annotating biological dark matter, as highlighted in a recent "Breaking the Bottleneck of Genomes" workshop report. It will also serve a national need for BER, Department of Defense, National Institutes of Health, and other federal agencies in the times of biological threat to characterize emerging pathogens to aid in the rapid development of new diagnostics and therapeutics.

Establishing the BioSCF will require investment in physical infrastructure to support large scale protein production, mass spectrometry characterization, high throughput biochemistry, and cellular imaging. In addition, new technologies will need to be developed, including flexible and scalable cell-based and cell-free transcription/translation expression platforms, high throughput microfluidics nanostructure–mass spectrometry-based assays of protein activities and binding, and new computational algorithms to extract information from large scale biochemical experiments.

2-year goal

Establish a distributed biological characterization facility that leverages existing Berkeley Lab resources.

Success metrics

- Enhanced and scalable pipelines for comparative viral genomic analyses developed and in place
- Ability to screen thousands of protein–small molecule interactions in a single day and demonstrate the capability to rapidly discover activities of proteins of unknown function
- SARS-CoV-2 proteins routinely produced and purified at scale under current good manufacturing practices (cGMP) for potential clinical studies
- SARS-CoV-2 proteins characterized for use in diagnostics or therapeutics
- Identified function for SARS-CoV-2 proteins of unknown functions, with an initial focus on ORF8, involved in host immune suppression and the severity of COVID-19 progression, as a proof of concept.

High-Throughput Biochemistry

A major limitation to achieving predictive biology is the lack of understanding of protein functions. Most functional assignments are based on homology to a relatively small number of proteins that have been experimentally characterized. This is often erroneous and completely fails for the large number of proteins that lack homology to characterized proteins. An opportunity now exists to integrate existing Berkeley Lab capabilities to create an integrated pipeline that enables rapid screening of tens of thousands of protein–small molecule reactions in a single day using droplet microfluidics integrated with nanostructure-initiator mass spectrometry (**Figure 2**) with liquid chromatography with tandem mass spectrometry capabilities that enable detailed protein characterization. Together this will enable rapid discovery of protein activities and detailed characterization of substrate specificities and kinetics.

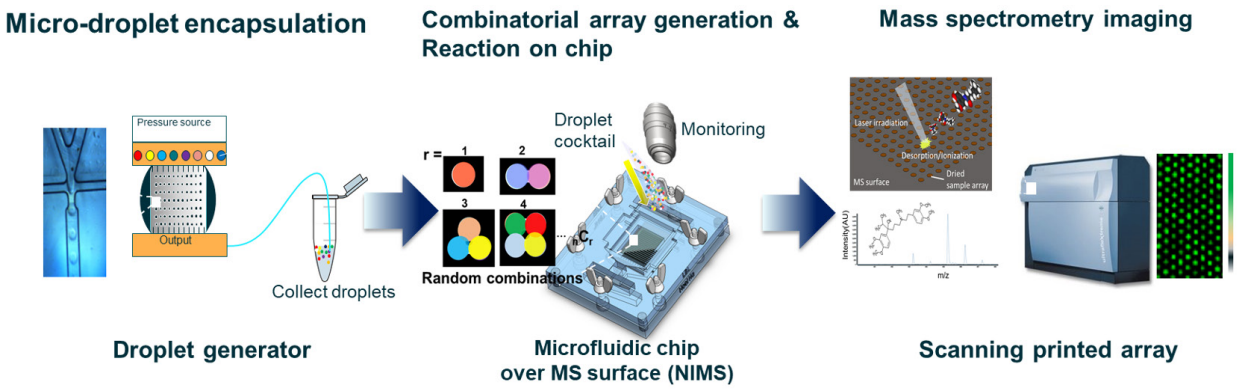


Figure 2: Droplet microfluidics integrated with nanostructure-initiator mass spectrometry enables screening tens of thousands of protein metabolite reactions per day.

Applying platform technologies for DOE mission research

DOE missions in biology will be greatly enhanced from a predictive understanding of activities of gene products, both proteins and nucleic acids. The Joint Genome Institute (JGI) continues to generate unprecedented numbers of viral, microbial, algal and plant genomes. However, the experimentally verified annotation of the gene products still lags behind the determination of the DNA sequence of genes. The BioSCF will enable the JGI to rapidly assign validated functions to important energy and environmental related proteins. Berkeley Lab projects such as Ecosystems and Networks Integrated with Genes and Molecular Assemblies (ENIGMA) are routinely isolating new environmentally important organisms from the field and have a need to determine their phenotype and it links to their genotype. These efforts, and those of many other research efforts, will be greatly enhanced by the BioSCF. The Joint BioEnergy Institute (JBEI) conducts extensive prospecting for new genes in environmental isolates to discover new lignocellulosic active enzymes. The BioSCF will greatly increase the speed of these new discoveries. Researchers at the Advanced Biofuels and Bioproducts Process Development Unit (ABPDU) are conversant in scaling the production and purification of a variety of proteins and have done so recently for three different companies pursuing COVID-19 solutions. Through the development of current Good Manufacturing Capabilities, they will be able to support the protein characterization efforts at Berkeley Lab and help scale-up for industry such that clinical trials will be possible from the samples generated at LBNL. The Agile BioFoundry (ABF) offers the ability to rapidly generate organisms through its design-build-test-learn cycle that leverages broad capabilities including machine learning, targeted proteomics, and microbial host domestication. The BioSCF will be an invaluable Berkeley Lab asset to help meet the goals that depend on harnessing plant and microbial metabolism for energy and the environment.

10-year goal

An integrated, brick-and-mortar biological characterization facility that is able to routinely and rapidly characterize thousands of gene products per year.

Success metrics

- Discovered novel small molecule–protein binding reactions of relevance
- Demonstrated ability to perform rapid and scalable comparative genomics of existing and emerging viruses.

- Identification of novel enzyme activities for proteins of unknown function from a wide array of organisms, including viruses, microbes and plants is routine
- Demonstrated data-driven and accurate prediction of gene function from diverse organisms, including those that have yet to be cultivated
- Developed a public platform for rapid protein production using novel promoter systems with tight repression and low-cost induction, and accelerated process development for scale-up
- Production of proteins with appropriate post-translational modifications via the application of agile, multi-host production infrastructure.

5-year milestones

- Demonstrated comprehensive biochemical, structural and computational characterization of thousands of novel gene products from hundreds of organisms annually
- Incorporated BioSCF capability through the FICUS program as a JGI user offering for internal and external use to support BER projects
- Automated inference of gene function from the integration of diverse data, including massive-scale genetics and biochemical data.

STRATEGY FOR INTEGRATION OF STRUCTURAL BIOLOGY AND IMAGING METHODS

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

COVID-19 and in general, pandemic-causing viruses, present extreme challenges. To return to activities without "COVID controls" (also known as non-pharmaceutical interventions) such as social distancing and face coverings, will require a multi-pronged approach of vaccines, treatments, and diagnostics. Development of these countermeasures requires a predictive and actionable understanding of each SARS-CoV-2 protein, its complexes, and its interaction with human host proteins. One of the powerful tools to gain such knowledge of the virus is visualization through atomic scale imaging, available through the field of structural biology. Berkeley Lab's infrastructure includes structural biology capabilities (X-ray and electron methods), imaging (IR, fluorescence, chemical mapping), computation/prediction, and protein-protein interaction mapping (X-ray footprinting, mass spectrometry). Some of these are well-established methods, such as macromolecular crystallography, while others are emerging methods, such as the Berkeley Lab Laser Accelerator (BELLA) Center imaging method based on compact plasma-based accelerators.

X-ray crystallography and cryo-Electron Microscopy (cryo-EM) provide high resolution information for deciphering protein mechanisms. Based on protein structures characterized by these methods, small molecules that could potentially block SARS-CoV-2 viral replication can be identified and subsequent high-resolution structures of the proteins with these inhibitors are necessary to improve the protein/inhibitor affinity for treatment of COVID-19 in patients. Structural analysis of antibodies with their antigens can aid in vaccine antigen design and perhaps predict the effectiveness of a vaccine. Berkeley Lab's Advanced Light Source (ALS) has eight well-established biological crystallography beamlines, and its beamline scientists are world experts in protein crystallography. Imaging methods based on compact plasma-based accelerators at the BELLA Center could increase biological crystallography facilities available for understanding the current pandemic.

Additional capabilities at the ALS are critical for informing an understanding of the mechanisms underlying infection with SARS-CoV-2 and COVID-19. Small Angle X-ray Scattering (SAXS) is a rapid means to quantitatively and structurally measure proteins and complexes in solution. It has been useful to determine how flexible SARS-CoV-2 proteins assemble. Its high throughput capability is powerful for screening for inhibitors that prevent protein-protein interaction or alter protein conformation. Another ALS facility provides X-ray foot printing, useful for mapping interfaces that can be used for understanding CoV-2 proteins or for antibody characterization. Understanding the mechanisms by which SARS-CoV-2 infects cells, replicates, and exits cells requires cellular imaging, ideally with as many modalities as possible.

Since COVID-19 reached pandemic level, the structural biology community has come together to determine the structures of every protein in the SARS-CoV-2 genome. These structures combined with biochemical analysis have revealed the complexity of viral proteins. The virus has a minimized RNA genome by which these viral proteins perform not one but multiple functions, so understanding each

protein's multifunctionality requires the determination of the underlying structural mechanisms. These include identification and analysis of the separate functions, and also how the protein converts from one function to another. Hybrid methods analysis, with inputs from multiple imaging modalities, will accelerate determination of SARS-CoV-2 protein mechanisms. While current research has expanded rapidly in response to the urgency of the pandemic, hybrid approaches and analyses are only just beginning to take shape.

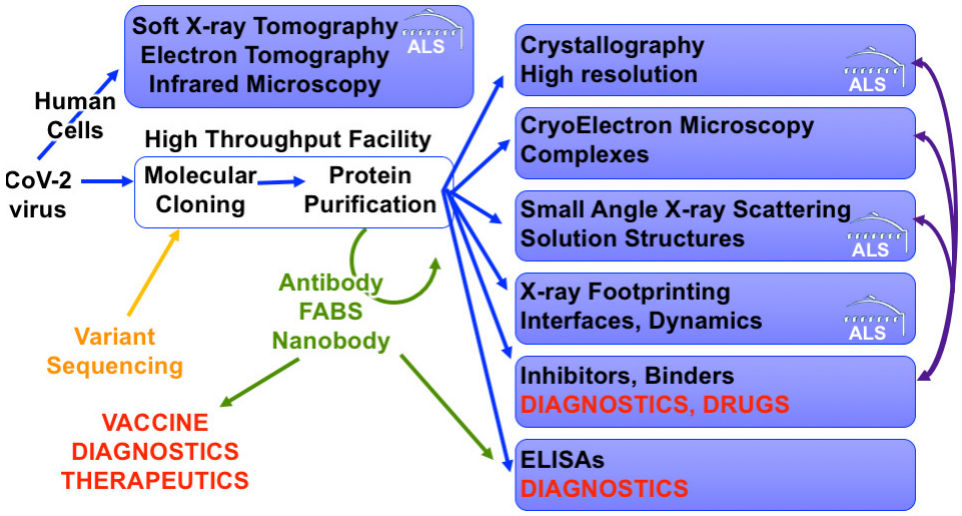


Figure 3: For accelerated development of effective SARS-CoV-2 vaccines, diagnostics, and therapeutics, Berkeley Lab Imaging Facilities coordinated with protein production facilities to enable a predictive understanding of viral mechanisms from atomic structures, interfaces, and inhibitors.

2-year goal

Provide proof-of-principle analysis of integrated structural biology methods with electron and photon imaging of biological material related to SARS-CoV-2 and COVID-19.

Success metrics

- Established a coordinating function to identify protein targets and combination of targets, integrate the different structural and imaging modalities as needed, and guide data analysis and dissemination
- Demonstrated capability to provide a single-shot image of a droplet and virus particle in a sample that is free-moving, not frozen or fixed in place
- Interrogated protein interactions in solution using two or more integrated modalities.

Single-Shot Imaging of Droplets

Berkeley Lab has been working on radiation sources that deliver flux in time-structured pulses, ranging from photons (UV, X-rays, and gamma rays) to particle beams (few-MeV and GeV-class electron bunches, as well as proton and ion beams). These bursts of short-duration, high-quality radiation emission are uniquely applicable to single-shot imaging, recording of dynamics, and/or reduction of exposure time to the sample, especially of interest if combined with time-synchronized secondary pulses for pump-probe setups. For example, droplet and macro-particle imaging in conjunction with laser-induced mass spectroscopy, or radiation-pulse-driven chemical tagging in conjunction with laser-induced fluorescence diagnostics.

Applying platform technologies for DOE mission research

Rapidly understanding complex biological interactions, for a pandemic or for environmental biology, requires a multi-pronged approach. Maximizing samples that can be tested by distinct modalities will provide a more comprehensive view that is necessary for understanding the complexity underlying biological mechanisms. Over the past decade, the SIBYLS beamline has forged new methodologies for combining SAXS with crystallography and cryo-EM to enable atomic resolution models enhanced with inherent dynamic features through computation. Integrated work at the Molecular Foundry and ALS has succeeded in bringing in X-ray footprinting and soft X-ray tomography (SXT). SXT and other microscopy methods can bridge from cells to tissues. Femtosecond X-rays with a high degree of spatial coherence find applications in unique imaging concepts, but also serve a diagnostic function for accelerator, allowing scientists to validate their understanding of injection physics, accelerator dynamics, and photon emission processes. Enabling single-shot imaging of single macro-molecules is intrinsically connected to accelerator beam physics towards the next generation of injectors and accelerator modules for TeV colliders.

10-year goal

A facility at Berkeley in which structural biology and imaging modalities are integrated and workflows for hybrid approaches have been developed, including data processing, analysis, and dissemination.

Success metrics

- Established pipelines from protein purification to different imaging modalities: crystallography, SAXS, cryoEM, and X-ray footprinting
- Demonstrated automated data analysis for the individual imaging modalities
- Developed hybrid methods for combining different imaging modalities.

5-year milestones

- Demonstrated fabrication and use of universal sample holders for two or more imaging modalities
- Development of infrastructure for locating precise regions of interest in samples as they are imaged using different modalities.

STRATEGY FOR HIGH-THROUGHPUT MATERIALS DISCOVERY AND CHARACTERIZATION

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

All around us, materials shape the ways in which humans interact with the world. Materials like wood and metal have been in use for millennia, while more recent innovations like semiconductors and plastics have been developed in the recent past. While many materials can be successfully employed to combat the SARS-CoV-2 pandemic, new and novel materials help to prevent the spread of the virus and treat COVID-19. For example, new materials or coatings for ventilation system filters killing the virus at contact would enable safe return to many indoor activities. To realize rapid discovery and generation of new materials for the pandemic requires the coordination of infrastructure — both instrumentation and coordinated expertise — in an efficient refinement cycle of design, synthesis, and characterization. Berkeley Lab already has much of this infrastructure in place in the form of synthesis instruments, years of collective expertise, and some of the nation's best characterization tools at the national user facilities.

These extensive capabilities and expertise can be leveraged through the creation of a cohesive and high-throughput center for the intelligent engineering of materials. This type of center will enable the highly efficient design and synthesis of new materials for diagnostics and mitigation strategies against SARS-CoV-2, and will also be able to rapidly pivot as needed to address emerging crises. For example, in the area of biomaterials, efforts are already underway to develop simple field assays to enable real-time detection of SARS-CoV-2 virus, while research in biopolymer design has led to novel antiviral therapeutics which are now in the animal testing stage.

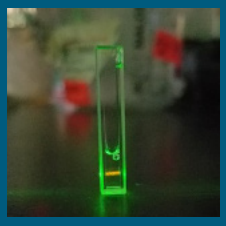
2-year goal

Establishment of an online project coordination portal and database that makes use of metadata structure to include diverse forms of data, including synthesis protocols, images, and scattering/diffraction data.

Success metrics

- Established a high-throughput cloning, expression and purification instrument suite for protein engineering
- Established specific rapid materials discovery approaches for high-throughput synthesis of peptides, peptoids, and nanoparticles targeted at COVID-19 applications
- Successful demonstration of the materials pipeline by development of a specific biomaterial for defense against SARS-CoV-2, such as an optical point-of-contact sensor capable of detecting 1 pM viral concentration.

Development of a Colorimetric Detection Method for SARS-CoV-2 Viral Presence



New methods are needed for simple detection of contamination of surfaces with SARS-CoV-2. As part of the NVBL initial phases, instant colorimetric assays for the presence of SARS-CoV-2 on masks were developed at Berkeley Lab. Detection is by eye and requires only a green laser pointer. In the image, a solution of SARS-CoV-2-binding peptides emits red light (mixed with green of the laser) when the viral receptor binding domain is added to the solution.

Applying platform technologies to DOE mission research

The diverse impact that can be made by novel materials is enormous and extends beyond the current pandemic to include applications such as self-repairing infrastructure materials, anti-corrosion smart coatings, selective membranes for water remediation and purification, catalysts for chemical production, novel drug-delivery mechanisms, and sensors for threat-detection. In the area of biomaterials, recent progress is based on extraordinary advances in methods for both manipulating and characterizing biological systems, such as the CRISPR-Cas system for gene editing, high-resolution cryo-EM for sub-cellular characterization, and high-throughput genome sequencing. Beyond biomaterials, in a separate but related activity, Berkeley Lab is actively planning the construction of an Advanced Materials Discovery Building (AMDB) which would enable similar advances in materials that are not necessarily biological in origin. The AMDB, in conjunction with a parallel biomaterials center, will bring together scientists and engineers in the diverse fields of physics, chemistry, biology and computing for the common goal of accelerated materials discovery.

Establishment of a center will leverage these new methods, as well as state-of-the-art instrumentation and world-leading expertise of its scientists to create a platform for deep, effective collaboration in a fast feedback cycle of synthesis, characterization, testing, and further refinement. This in turn enables teams to respond quickly to emergent threats, beyond even the COVID-19 pandemic.

10-year goal

Establishment of a coordinated center for materials discovery.

Success metrics

- Established effective teams of scientists for rapid materials design and characterization, making use of a materials discovery portal and high-throughput instrumentation
- Designed and synthesized materials capable of rapid field detection and quantification of pathogens and toxins in small quantities
- Demonstrated synthesis of a library of building block (“plug and play”) materials for construction of more complex structures for energy capture, storage, and transfer
- Synthesized bio-compatible delivery materials for drugs, imaging and monitoring devices into humans and other higher order organisms

- Designed and synthesized successful practical materials, such as an indicator material for the rapid detection of SARS-CoV-2 virus particles in the environment, or a peptoid delivery system for a novel CRISPR system that directly targets the viral genome
- Demonstrated complete rapid response (genomics, structure, materials) to newly evolved pathogens or manmade toxins.

5-year milestones

- Established an integrated lab wide data platform for diverse experimental data, accessible via the project coordination portal, and making use of metadata structures
- Demonstrated an automated high-throughput sequence defined biopolymer materials discovery platform with AI/ML algorithmic design process feedback
- Established an automated high-throughput protein engineering workflow
- Demonstrated a successful genome engineering project to program an organism to produce a biomaterial with a given characteristic, such as permeability, conductivity, or hardness.

STRATEGY FOR MICROSYSTEMS, ROBOTICS, AUTOMATION, AND SELF-DRIVING LABS

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

One of the challenges for combatting the SARS-CoV-2 virus and the COVID-19 pandemic is the current lack of technologies for high throughput testing for the presence of the virus. Research is needed to drive testing for humans (though sampling nasal oropharyngeal swabs, saliva, blood, organs and even breath), animals, facilities (both air and surface sampling), and in ecosystems (including wastewater). Additionally, closed loop discovery engines with a rapid engineering cycle for therapeutics and diagnostics could enable swift development of antiviral compounds. Microsystems, such as digital microfluidic devices, will be critical to miniaturizing and automating experimentation, along with new automation for data processing for challenging formats like those created through different imaging technologies. "Self-driving labs," where a large training dataset (thousands or millions of observations) for computation can lead to the automated design of follow-on experiments and well-defined criteria for success, will free humans from repetitive tasks that instrumentation can do and instead allow them to focus on important knowledge work to enable the rapid discovery and prediction for new COVID-19 treatments. Linking these elements in a test bed to test new robotics, high-throughput technologies, and automated AI-driven experimentation will enable the prototyping and evaluation of remote experimentation techniques. These advances have the dual benefits of enabling faster scientific discovery, and reducing exposure of humans to potential workplace risk during pandemic conditions.

Automating workflows through linking robotic unit operations, including sample preparation, manipulation, and observation, can speed the time to discovery for new anti-viral medications and treatments for COVID-19. Additionally, these automated workflows can allow for robust testing procedures with minimal staffing and reduced human error. The Berkeley Lab Automated Diagnostics Experimental Lab (BLADE) was established in 2020 to develop and evaluate new automated and high-throughput testing procedures for the detection of SARS-CoV-2. The introduction of microsystems, or miniaturized experimental equipment such as digital or droplet microfluidic devices, can result in reductions in the quantities of reagents needed and accelerate experimentation and diagnostics, further enhancing the ability to detect the virus and establishing methods that could be deployed outside of clinical lab settings, such as a doctor's office or urgent care facility. Finally, through the application of AI and machine learning to the large datasets enabled by automated experimentation and microsystems, new experiments can be recommended and implemented without human intervention. These self-driving labs will be able to quickly drive to potential treatments and diagnostic procedures, while also allowing researchers to work remotely during pandemics or other crises where access to research facilities is limited.

2-year goal

Demonstrate an intuitive, deployable microsystem for multiplexed antibody and SARS-CoV-2 detection that can be used as laboratory or point-of-care diagnostic technologies with minimal training.

Success metrics

- Demonstrated an easy to operate deployable detection device has been demonstrated to monitor facilities (e.g., surfaces/airflows) and human/animal populations with results in minutes to hours
- Designed, developed, demonstrated and deployed microsystems that can be used as laboratory or point-of-care diagnostics for SARS-CoV-2 and other targeted respiratory viruses that are considered significant public health risks
- Developed deployable detection devices further applied to monitoring viral residence and fate within fabricated ecosystems
- Completed a published report documenting interchangeable devices, reagents, workflows, materials in the context of antibody and SARS-CoV-2 detection that has reduced the number of single-source dependencies subject to unreliable supply chains
- Demonstrated ability to pool samples and barcoding of samples for unique sample tracking and readouts
- Demonstrated refinement of an automated experiment through artificial intelligence/machine learning.

Deployable and Miniaturized COVID-19 Diagnostic Equipment

Microsystems have the potential to reduce processing time and reagent usage. With Sandia National Laboratories, Berkeley Lab has developed a digital microfluidic platform that is capable of real-time polymerase chain reaction (RT-PCR) in micron-sized droplets. Microfluidic devices offer small thermal mass, low thermal inertia, and rapid heat transfer. The small volumes involved significantly reduce sample and reagent consumption, leading to inexpensive operation of the systems. However, in most cases volumes are still large enough for advantageous bulk kinetics to apply. Through further development of this system, future studies are aimed to amplify COVID-19 RNA as a miniaturized RT-qPCR workflow similar to conventional detection workflows for assays approved by World Health Organization and Centers for Disease Control. The ultimate goal is to engineer these microsystems into ruggedized platforms capable of being operated in a standalone fashion in laboratory and clinical settings, including their use in point-of-care environments.

Applying platform technologies for DOE mission research

Self-driving labs would enable human experts to work on challenging high-level problems: deciding on the important questions, designing the experiments, and consolidating the obtained knowledge. These changes would accelerate scientific progress by orders of magnitude. Self-driving labs provide an interactive environment in which AIs can actively learn from the consequences of their predictions. This paradigm is closer to how human intelligence learns: not so much from labeled examples, but rather from interacting with and probing the “environment.” Significant developments in disciplines such as Reinforcement Learning or control theory are needed, as are new disciplines centered on extracting actionable knowledge from AI agents.

In concert with AI, microsystems and automation of workflows will enable rapid generation of data and improvement of experimentation. To fully realize automated experimentation and self-driving labs, several laboratory unit operations will need to be integrated and automated, with impedances and interdependencies managed. From genome sequencing and protein production to materials and chemicals design and production, automated and robotic technologies are becoming commonplace. Through miniaturization of bench top (or larger) equipment to microsystems, reagents can be conserved, and experiments accelerated. Efforts such as the Agile BioFoundry, a Berkeley Lab led consortium of National Laboratories, the nascent effort to develop robotics to manage fabrication ecosystem experimentation, and high-throughput protein crystallography pipelines can all be enhanced through further automated experimentation. A test bed that provides interchangeable unit operations for integration, testing, and evaluation will speed the development of robust workflows for new microsystems, automation, and robotics.

10-year goal

Establish a pre-competitive test bed that enhances the commercialization, resilience, and impact of automated experimentation and self-driving labs for the research community.

Success metrics

- Demonstrated an integrated robotics, automation, and AI/machine learning platform that can refine experimental design without human intervention
- Demonstrated that operations have been made more resilient and efficient when faced with periods of fluctuating staffing levels and instrument downtimes
- Robotics/automation capabilities gainfully applied to enable/accelerate other strategic scientific research areas
- An industrial consortium has been established to support a pre-competitive test bed for automated unit operation swapping
- Contextually de-risked automated unit operations have been more extensively commercialized and/or deployed across facilities/organizations
- Deployed microsystems that can be used in the study of ecosystems, perform all aspects of the design-build-test cycle of biomanufacturing, measure he health of organisms in bioreactors, or detect a wide range of pathogens, select agents, and toxins known to pose significant risks to human health.

5-year milestones

- Closed loop discovery engine applied to two distinct scientific or technological research and development areas
- Microsystems originally developed for SARS-CoV-2 detection adapted/extended to at least one additional application area such as biodefense or emerging infectious diseases
- Automated laboratory fabricated ecosystem capabilities provided key insights into the genetic and ecological processes controlling microbial transmission, persistence, and fate
- Developed and demonstrated a digital microfluidics system that can accurately monitor rhizospheres in terms of metabolites known to be critical components of central and secondary metabolism
- Applied self-driving lab capability to robotic fabricated ecosystem experimentation.

STRATEGIES TO ADVANCE SARS-CoV-2/COVID-19 AND DOE MISSION RESEARCH

STRATEGY FOR INVERSE DESIGN OF COMPOUNDS

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

Currently, there are no fully validated effective antivirals against the novel coronavirus, SARS-CoV-2. Vast chemical libraries consisting of millions of compounds have been screened by drug companies and public efforts within the DOE complex. While candidate antivirals have emerged, the explored space of around six million compounds is infinitesimal compared to the combinatorially vast space of realizable drug-like compounds. There are at least 10^{60} druglike compounds that could, in principle, be chemically or biologically synthesized. To identify useful compounds, Berkeley Lab proposes to solve the inverse problem and transform drug discovery into a tractable optimization procedure.

The inverse design of compounds has long been a moonshot goal of drug companies, materials scientists, polymer chemists, and engineers. However, only recently have generative models and compound AI architectures placed this goal in reach. To design an antiviral, molecules that are predicted to interact with target proteins in specific ways need to be identified. Ideally, such design activities would constitute massively multi-objective optimization to simultaneously ensure antiviral activity while ensuring low toxicity in humans along with bioavailability.

2-year goal

Using an inverse design approach, demonstrate antiviral activity against SARS-CoV-2, the chemistry and synthetic biology required to produce those designed antivirals, and develop controls for industrial-scale bioreactors and/or chemistry platforms for production.

Success metrics

- Optimized compounds that significantly outperform known, currently available compounds, along with methods to produce these compounds
- Demonstrated compound production at a scale necessary for evaluation
- Controlled and executed compound production at a scale relevant for technology transfer to biomedical production facilities.

Applying this research to DOE mission science

Combining generative models with reinforcement learning and multi-objective optimization constitutes a general framework for inverse design. To design antivirals, the *in silico* nature of molecular docking simulations can be integrated with AI — prioritizing important "measurements" via simulation, learning from these, and iterating until optimal (or sufficiently well-engineered) compounds are generated. This procedure is central to the mission space of DOE's Office of Advanced Scientific Computing Research and is made possible by access to high performance computing (HPC). Others have integrated AI hypervisors with beamline protocols to optimize the utility of physical measurements. Hence, the framework includes, but transcends, AI-simulation integration.

Optimizing biofuel design and production through integration with the design of chassis organisms and combustion simulations is an obvious application of AI-enabled inverse design. Going further, coupling organism design to the design and control of "smart" bioreactors is similarly realizable. These steps require close coordination between research programs to develop the training data and iteration/learning cycle required for machine learning. The same argument can be made for design and synthesis of biopolymers and/or genome-encoded materials.

The design of bespoke chemicals, compounds, polymers, and more broadly materials through goal-oriented inverse design and multi-objective optimization is within reach. What is needed is a strong push in these directions to transform disciplines including synthetic biology, polymer chemistry, the design of materials, and the means of their production in the emerging bioeconomy.

Natural Language Processing and Machine Learning to Facilitate Discovery Science, Data Integration and Dissemination

Natural Language Processing techniques can capture 'latent' knowledge in scientific text and be used to draw novel connections that had been overlooked by human researchers. Similar techniques could be applied to the COVID-19 research literature, for tasks such as drug re-purposing, identifying gene/disease links, and highlighting knowledge gaps about the virus. In particular, adapted Bidirectional Encoder Representations from Transformers (BERT) methods will be used to identify associations that are not obvious in text that can advance an ability to treat and mitigate SARS-CoV-2 and COVID-19.

10-year goal

Develop an AI platform for the inverse design of compounds and materials for targeted applications in energy, environment, and biomedical research

Success metrics

- Developed optimized compounds with improved performance over incumbents that can serve as components in drop-in biofuels
- Produced new materials and their precursors based on target properties for specific energy applications, for example window coatings
- Produced a desired compound at an industrially relevant scale using predictive tools and controlled reactors.

5-year milestones

- Developed user-friendly tools for compound prediction based on specific end-use applications that are available to researchers
- Demonstrated that tools for Natural Language Processing and machine learning that can recommend integrated production pathways for compounds of interest based on available literature
- Developed a comprehensive understanding of how to scale new compounds to industrially relevant quantities.

STRATEGY FOR ECOSYSTEM SURVEILLANCE

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

Transmission of SARS-CoV-2 and mortality rates of COVID-19 have differed profoundly across the continental United States. Light intensity, temperature, and humidity have been identified as important factors in viral transport and activity, and hence, transmission. Population demographics, mobility data, and a variety of surrogate measurements for human behavior are explanatory factors for transmission rates. In contrast, factors predictive of disease severity include comorbidities, genetics, and pre-existing cross-immunity. However, no sufficiently comprehensive study as yet exists to enable the assessment of the relative strengths of these effects, or how much variance they explain, to contribute to a spatially resolved risk of infection. Knowing both the risk of infection and the risk from infection would be valuable for selectively resuming activities paused due to the pandemic and to understand how to prioritize vaccination efforts once vaccine are available.

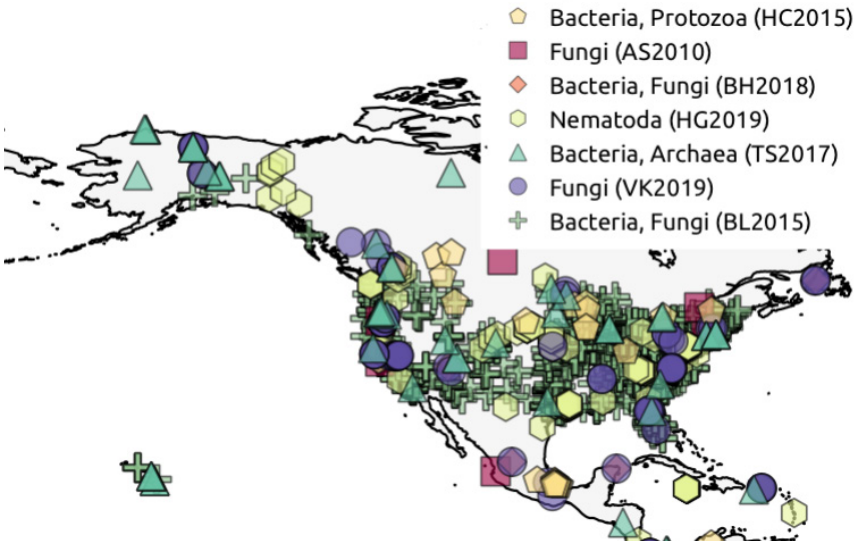


Figure 4: A map of published microbiome surveys of the US.

Ecosystem surveillance may be able to mitigate challenges posed by the current COVID-19 pandemic and to prevent future pandemics before they begin. Environmental metagenomic sampling can be used to enhance the resolution of risk indicators using, for example, indoor microbiomes as information-rich sensors that correlate to virus transmission and disease progression. Sampling of points of urban and ecosystem convergence such as waste and surface waters, or indoor surface microbiomes, has the potential to provide early indicators when as few as one in 10,000 individuals are infected. By selectively monitoring specific facilities, followed by individual testing and contact tracing in the event of a positive test, such tools have the potential to significantly slow, or even halt, the spread of the virus.

Examples of successful ecosystem monitoring efforts include the surveillance of livestock and wild animal populations, particularly by the National Wildlife Research Center, to identify avian and swine populations at risk of propagating flu strains to humans. In fact, modest investments in surveillance and culling activities have successful prevented major zoonotic transmissions from these reservoirs in the US for nearly a decade. Similarly, Centers for Disease Control programs monitoring influenza risk have facilitated the prioritization of resources (county-by-county) to combat the economic impacts of seasonal flus.

Widespread sequencing analysis of SARS-CoV-2 during the pandemic has uncovered a host of signatures of adaptation to human populations affecting efficient transmission. Studies have revealed significant impacts of environmental factors, including those of the environmental microbiome on human susceptibility to disease, and may indicate such a role for environmental factors for modulating the transmission and severity of COVID-19 regionally across the US.

2-year goal

Determine the local factors that modulate transmission rates and disease severity, including climate, environmental microbiomes, human genetics, behavior and socioeconomic traits, to develop metrics of intrinsic hazard for specific populations and geographic regions that take into account past environmental and microbiome exposures.

Success metrics

- Demonstrated accurate prediction of localized mortality rates upon SARS-CoV-2 infection for specific regions of the United States based on specific environmental factors, including microbiomes and their constituents
- Demonstrated accurate prediction of localized transmission rates of SARS-CoV-2 in specific regions based on key environmental factors identified through environmental surveillance
- Identified factors enabling accurate prediction of mortality rates and infection doubling times and their relationship to observed microbiomes in specific regions.

An Ecosystem Observatory for COVID-19

For the first time, computational power now exists to model the co-evolution of species and viruses in natural ecosystems to predict the emergence of zoonotic hops before they occur. A distributed observatory, coupled to high-performance computing, can detect and predict the trajectories of ecosystem dynamics and identify transitions in metastable equilibria before they occur. Further, the identification of environmental and microbiome factors that mediate disease severity can aid in the prioritization of resources and public policy interventions to minimize total harm from current and future pandemics.

Applying this research to DOE mission science

Patterns in global travel and trade, changing land use, pressures from changing climate regimes, urbanization, and population growth are accelerating the expansion of transmission and habitat range of numerous disease vectors of ecosystem, economic and human health importance. New approaches to surveillance have been called for that encompass tracking of both pathogen and habitat dynamics. Emergence of new human pathogens, for example, is preceded by long periods of adaptation to vertebrate species, and in retrospect, are often “predictable.” This pathogen-host evolutionary struggle

also plays out in Earth's ecosystems, and such microbial invasions also threaten the availability of clean water and the production of food, fuel, and fiber. Currently, observations of environmental stressors and pathogen evolution occur independently, and signs of pathogen invasions are recognized only after widespread infection and damage.

Ecosystem surveillance, including (meta)genome sequencing campaigns, multi-omics approaches to understand gene function, and new platform technologies for local and remote ecosystem sensing, can reveal deep insights into the coevolution of species in ecological and other contexts.

10-year goal

Establish an integrative ecosystem surveillance capability and the data processing required to identify the current states of species interactions in dynamic and heterogeneous landscapes, and to predict future states of interaction based on predictions of future states of climate.

Success metrics

- Demonstrated accurate early-warning systems for detection of microbial invasions in natural and managed ecosystems
- Accurately predicted pathogen and vector clines and their environmental constraints relevant to bioenergy feedstocks and natural ecosystems
- Developed a predictive understanding of ecosystem services and the risks associated with changing clines in natural and managed ecosystems.

5-year milestones

- Developed and deployed of a multi-parameter artificial intelligence-based surveillance system for real-time ecosystem monitoring for microbial invasions and other perturbations
- Developed *In silico* and reduced order (fabricated ecosystems such as EcoPODs, SMARTSoils) models of ecosystem response to microbial invasion
- Developed a predictive model of molecular microbial pathogen coevolution at ecosystem scales
- Integrated physical and biological remote sensing data with molecular surveys and developed surrogate models for novel adaptation events at landscape scales.

STRATEGY FOR UNDERSTANDING THE ROLE OF HUMAN MOBILITY IN THE SPREAD OF COVID-19

Having been declared as a global pandemic by the World Health Organization (WHO), COVID-19 poses a significant threat to public health and the global economy at large. Although much is still not known regarding the spread of COVID-19 within the human population, substantive evidence indicates that this virus is transmitted from person to person by direct contact and airborne transmission of the virus via emitted respiratory fluid particles. Under these circumstances, detecting the origin and emergence of new disease infections such as COVID-19 is critical to preventing and reducing their spread, particularly in early stages before they become a pandemic. A capability to track the geographic spread of COVID-19 cases provides useful information for disease prevention and control. One of the ways diseases (including COVID-19) spread is through the continuous spatial exchange of susceptible, latent and infectious individuals by travel. Thus, it is important to examine the role of human mobility in the geographic spread of COVID-19 cases within a large metropolitan area. This can be done by leveraging high performance computing (HPC) based agent-based modeling (ABM) that simulates regional travel patterns, including activity engagement, activity timing, mode choice, and social networks, in both personal and shared (transit, ride-hail, and micro-mobility) vehicles. Further, the agent-based human mobility model could be coupled with epidemiological models to better understand and investigate the role that transportation systems play in the spread of COVID-19 among the population. Research to design and develop a higher-order agent-based simulation model focused on understanding the role of human mobility and air travel in the spread of COVID-19 disease within a large metropolitan area would support such a goal. Berkeley Lab could leverage the current ongoing efforts under the COVID-19 rapid response modeling funded by DOE's NVBL program, in coordination with DOE's Energy Efficiency and Renewable Energy Vehicle Technologies Office (VTO).

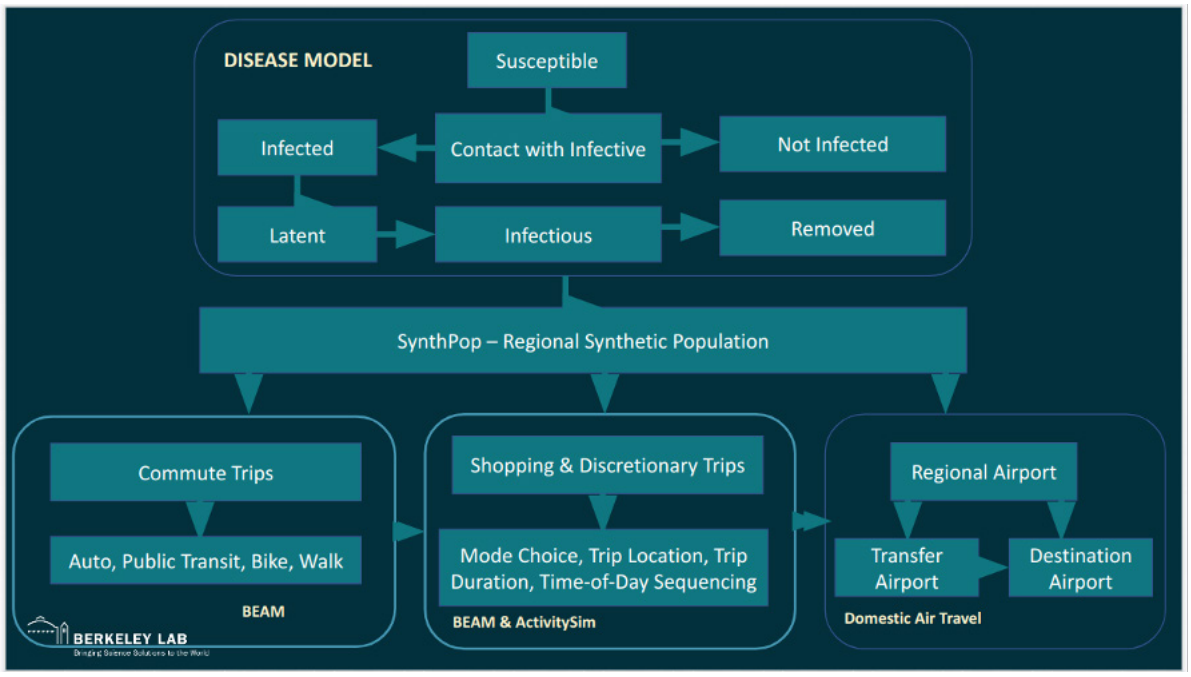


Figure 5: Integrated Epidemiology and Mobility Systems Analysis.

Addressing scientific challenges posed by SARS-CoV-2 and COVID-19

2-year goal

Deliver a calibrated and validated integrated model of population mobility and disease transmission, leveraging the ABM framework and HPC resources at LBNL in order to provide a tool for policy makers for informed decision-making.

Success metrics

- Systematically identified and integrated the interconnectedness of population mobility with disease transmission into the agent-based models
- Comprehensively calibrated and validated both the population mobility (transportation system) model and disease transmission model in New York City that capture the various phases of COVID-19 outbreak from the onset of the pandemic through the reopening phases
- Engaged with stakeholders to develop scenarios to evaluate policies to mitigate COVID-19 spread that also achieve safe and efficient movement of the population within a region
- Provided a tool to conduct contact tracing, physical interactions, and use of time on activities (both indoors and outdoors) of suspected infectious individuals to public health experts, in order to design and evaluate various policy interventions to minimize virus spread and protect public health.

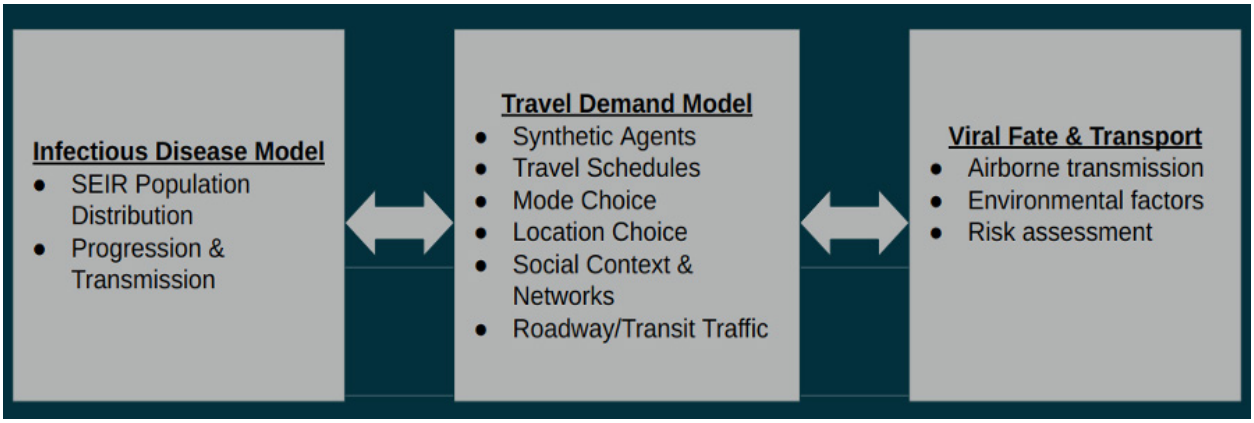


Figure 6: Integrated Systems of Mobility and Disease Spread and Viral Transport.

Applying this research to DOE mission science

Improved human mobility and transportation system modeling, which can be adapted to evaluate responses to pandemic and other large-scale emergencies, can also be applied to understand the efficient and effective use of various modes of transportation. An opportunity now exists to integrate existing models related to epidemiology, population movements and airborne transmission into a single modeling system to develop a comprehensive picture of the larger system-wide impacts for any given metropolitan area. The integrated systems approach will leverage HPC resources at DOE facilities, which are necessary to quickly run large scale simulations of millions of agents in metropolitan regions.

Rapid Prototyping and Deployment of Mobility Systems Analysis in New York City

A major limitation of existing transportation simulation models is that resources (financial, data, and computational) are required to represent the detailed behavior and complex interactions of agents in an urban environment. As a result, many public agencies shy away from investing resources in developing or using these tools, as they require a lot of data and maintenance in order to conduct evaluations of various policies. However, the current on-going DOE NVBL project, using the BEAM model in the New York City metro region, serves as a platform that policy makers and transportation modelers can rapidly prototype a localized model using publicly available data. The agent-based transportation model BEAM (which mimics traveler behavior of a regional population) could be integrated with traditional epidemiological models (e.g. standard compartmental model) to explicitly recognize the role of detailed traveler behavior in the spread of disease throughout the population. Needed is a higher-order agent-based simulation model that takes into account individuals' movements throughout the metropolitan area as well as inter-regional air travel (and/or long-distance travel) to understand disease transmission across a larger population in the U.S. to further enhance the understanding of the spread of COVID-19.

Finally, the proposed research strategy also aligns with other ongoing efforts within the recently created VTO Systems and Modeling for Accelerated Research in Transportation (SMART) program to improve these mobility modeling tools and make them accessible and reproducible to multiple regions in the U.S. Through a wider systems analysis, a combination of scenarios and relevant human behavior on systemwide mobility, traffic congestion, and energy consumption in a region can be examined.

10-year goal

Develop actionable interventions for transportation in metropolitan areas from integrated modeling that account for mode choice, human behavior, and emergency response needs.

Success metrics

- Completed integration of mobility, disease transmission and viral fate transport models that are calibrated and validated for several major metropolitan areas in the U.S.
- Stakeholders engaged for targeted responses in both policy development and evaluation aimed towards various outcomes in emergency response and disaster preparedness
- Developed a viable user-base that leverages the open source nature of Berkeley Lab's Behavior, Energy, Autonomy and Mobility transportation simulation platform to further design, develop and deploy integrated models for various planning purposes at state-, regional-, and city-level jurisdictions.

5-year milestones

- Successfully demonstrated the integrated modeling approach that takes into account the mobility, disease transmission and viral fate transport mechanisms for larger systems analysis
- Calibrated and validated the integrated model in a large metropolitan area (e.g. New York City) to demonstrate the capability and scale of advanced models for emergency responses
- Worked with various stakeholders nationwide for outreach and interest in this approach to be able to adopt and deploy the model for their localized needs in policy evaluation.

BERKELEY LAB ASSETS TO ADDRESS SARS-CoV-2 AND COVID-19

Berkeley Lab User Facilities

Advanced Light Source

The ALS provides access to macromolecular crystallography capabilities, as well as high-throughput small-angle x-ray scattering (SAXS) structural characterization and soft x-ray tomography. SAXS can also play a role in screening for interactions between viral proteins and potential diagnostic probes and also candidate therapeutics. Experimental data, like SAXS, can also be used as a powerful guide for computational modeling efforts. Soft x-ray tomography is being used to image organelles inside cells, allowing comparison between SARS-CoV-2-infected and uninfected cells. The ALS's infrared capabilities could be used to measure specific chemical signatures for viral detection.

DOE Joint Genome Institute

JGI performs large-scale sequencing of DNA and RNA and computational genomic analyses, and is now using these capabilities for comparative and phylogenomic analyses of coronavirus genomes. These analyses could be expanded to include coronaviruses from diverse host species and environments and SARS-CoV-2 genomes from diverse patient populations to investigate evolution, mutation frequency and constrained regions, epigenetics, and/or pathogenicity, as well as to identify potentially accessible regions of surface proteins for immuno-targeting. JGI capabilities also enable genomic analyses of host-virus interactions and host microbiome changes with viral infection.

Energy Sciences Network

ESnet's high-speed network, engineered and optimized to support the Department's large-scale scientific research, interconnects the entire national laboratory complex, including its supercomputer centers and user facilities. ESnet allows scientists to use DOE's unique research facilities independent of time and location with state-of-the-art performance levels by providing direct connections to more than 40 DOE sites at speeds up to 100 gigabits per second (Gbps).

Molecular Foundry

The Molecular Foundry is developing point-of-contact, instant colorimetric probes of SARS-CoV-2 contamination on masks and other surfaces. In addition, it has automated high-throughput biopolymer synthesis and characterization capabilities for the rapid discovery of potential therapeutic or diagnostic agents. It also has capabilities for developing specialized nanostructures that could form the basis for intrinsically antiviral surfaces and materials. A collaboration between the Molecular Foundry and the Materials Sciences Division has developed equipment for testing the particle filtration efficiency of masks and respirators.

National Energy Research Scientific Computing Center

NERSC provides world-class supercomputing and data resources to support simulation, data analysis, and AI/machine learning efforts across the Office of Science. NERSC also contributes the expert staff and expertise that enables researchers to reach the scale and performance required to take full advantage of

its systems. Computing and data analytics touch all areas of science and NERSC's resources provide a scale that is available at only a handful of other facilities. NERSC is participating in the COVID-19 HPC Consortium as a resource provider and supports a number of COVID-19 related research projects.

Collaboration Facilities and major research programs

The **Agile BioFoundry** integrates industrially relevant production microbes, advanced tools for biological engineering and data analysis, and robust, scaled-up processes for integrated biomanufacturing. These capabilities can be focused on the development of new therapeutics for SARS-CoV-2 and COVID-19.

The Advanced Biofuels and Bioproducts Development Unit (**ABPDU**) is a bioprocess scale-up facility that enables early stage innovation by performing process science to transfer fermentation and downstream purification technologies from microtiter plates to bioreactors at 100s of liters. By working with three industrial entities on COVID-19 projects, the ABPDU has demonstrated the need for such a capability to rapidly accelerate novel technologies towards commercialization during a pandemic.

The **BELLA Center** has high-power laser and particle accelerator capabilities that can be used for mass spectrometry and time resolved x-ray imaging to characterize viruses and their microenvironments in droplets. This can help researchers understand the survival of the virus in air and on surfaces.

The Berkeley Lab Automated Diagnostics Experimental Lab (**BLADE**) was established to develop, evaluate, and de-risk new methods and strategies for SARS-CoV-2 diagnostic testing. BLADE integrates liquid handling automation, molecular biology, genomics and sequencing, and analysis.

FLEXLAB lets users test energy-efficient building systems individually or as an integrated system under real-world conditions. FLEXLAB testbeds can monitor and assess heating, ventilation, air conditioning, lighting, windows, building envelopes, control systems and plug loads — in any combination — and is being used now for SARS-CoV-2 particle experiments.

The Joint BioEnergy Institute (**JBEI**) is a U.S. Department of Energy Bioenergy Research Center dedicated to developing advanced biofuels — liquid fuels derived from the solar energy stored in plant biomass that can replace gasoline, diesel and jet fuels. Researchers are using the latest tools in molecular biology, chemical engineering, computational and robotic technologies to transform biomass into biofuels and bioproducts. JBEI has developed high-throughput microfluidic systems that can be used for large scale screening of samples to help develop diagnostic probes.

Key Berkeley Lab capabilities

Electron cryo-microscopy (cryo-EM) and related modalities are available to examine the structures of key viral proteins and their interactions with candidate antibody diagnostics and therapeutics. Individual-particle electron tomography (IPET) and liquid-cell TEM can be used to visualize an individual virus infecting a cell in 3D with nm resolution.

cDNA library construction: The construction of cDNAs for each viral protein, and each mutation, is a prerequisite for structural biology and the development of antiviral agents. LBNL has the capacity to develop thousands of cDNA expression constructs, a resource likely to be of enduring value.

Computational algorithms and knowledge bases: A variety of tools are available using DOE's high-performance computing facilities to build high-resolution models of viral proteins and to use AI algorithms to computationally design antiviral compounds. These can also be used to create prioritized lists of viral and human targets and protein complexes, and perform network analysis of interacting viral and human genes, and evolutionary and mutational analysis of viral variability. Aggregating the multiple relevant data types, their context, metadata, and provenance, will be critical to supporting real-time

efforts for COVID-19 response. The DOE Systems Biology Knowledgebase has extensive experience in these areas, along with the National Microbiome Data Collaborative and the Gene Ontology group. Berkeley Lab's expertise in computational structural biology can also be used to create optimized models from crystallographic and cryo-EM data to guide the modeling of antiviral compounds.

Virus contagion and risk models: Expertise in indoor environmental air quality and contaminant transport can be used to develop improved quantitative contagion and risk models using data from various indoor environments (hospitals, cruise ships, classrooms, and others). This information can be used to improve estimates of transmissibility in various indoor environments to aid in developing preventative strategies.

Viral protein production, purification, and antibody development: Capabilities for the safe synthesis of viral proteins (in large quantities via the ABPDU), developing antibodies, assessing and optimizing the specificity of antibodies, and identifying monoclonal lines if possible, for production.

Assess individual susceptibility to COVID-19: Capabilities can be used to gather individual, clinical and/or genetic data, e.g., whole genome sequencing or targeted sequencing (MHC locus and others) and construct models of individual susceptibility including clinical and genetic factors.

Transgenic/engineered mouse production and line expansion: The Mammalian Functional Genomics group within the Environmental Genomics and Systems Biology division performs large-scale generation of transgenic and engineered mice. Equipment and trained staff are ready to perform pronuclear injection of CRISPR-Cas and DNA into fertilized mouse eggs, embryo transfers to surrogate animals, and in vitro fertilization to rapidly expand mouse lines. These capabilities could be harnessed to generate "humanized" mouse models for research studies of coronavirus infection.

High-resolution climate forecasting and re-analysis: A number of programs provide actionable and credible climate modeling products for specific regions. These products can be used to interpret and project environmental parameters at the city block scale that are key to virus transport and survival in both indoor and outdoor environments.

APPENDIX: RECOMMENDATIONS FOR COLLABORATION AND TEAM BUILDING

The instantiation of the National Virtual Biotechnology Laboratory and resulting DOE-funded research necessitated building new collaborations between National Laboratories and researchers in a very short period of time. Should any additional funding be available through NVBL or DOE, it is likely that new research projects will be executed in collaboration with others. At the visioning workshops, participants were asked to provide recommendations for collaborations with other national laboratories and for team building, based on experience from the rapid start of NVBL and from other similar efforts.

Collaborations

Participants noted that while new collaborative projects were quickly built for NVBL research, Berkeley Lab collaborates regularly with other National laboratories. Some of the new NVBL collaborations were the result of multiple labs offering expertise typically applied elsewhere for the multidisciplinary research themes to address the COVID-19 pandemic. The multidisciplinary nature of the research also brought together collaborators that typically do not work together, even within the same national lab. Participants recommended deepening collaborations with other national labs in the following areas, though this is not an exhaustive list of potential collaboration space or collaborating National Laboratories:

- Epidemiological modeling: Lawrence Livermore National Lab, Los Alamos National Lab
- Biosafety Level 3 laboratory facilities: Lawrence Livermore National Lab, Pacific Northwest National Lab, Los Alamos National Lab
- Protein structure and characterization: Argonne National Lab
- Viral fate and transport: Argonne National Lab, Sandia National Labs
- Functional genomics and understanding of virus and host responses: Pacific Northwest National Lab, Oak Ridge National Lab, Lawrence Livermore National Lab, Los Alamos National Lab, Sandia National Labs.

Additionally, participants suggested that partnership with universities, particularly those with medical schools, would provide critical clinical expertise for SARS-CoV-2 and COVID-19 research. University of California San Francisco was highlighted as an obvious and potentially strong partner based on the current relationship between Berkeley Lab and UCSF. Participants also discussed the importance of partnering with private sector companies to ensure that knowledge and technologies developed at Berkeley Lab are transferred to companies that can manufacture and scale new testing methods, therapeutics, or other products. Additionally, Berkeley Lab can serve as a key partner for companies to de-risk their technologies and to provide key insights into how products like new therapeutics prevent or treat COVID-19.

Team building

During the visioning session, participants were asked to provide some lessons learned from the NVBL teaming activities and some possible approaches to move forward. Generally, there was agreement that the short time period to assemble the NVBL research teams made it challenging to identify all the needed expertise for the research themes and to ensure an inclusive approach. In the future, participants provided some of the following ideas to facilitate establishing teams of researchers:

- Encourage continued importance of solving national-scale problems
- Use of surveys and/or Dear Colleague letters to surface concepts from unexpected sources
- Foster inclusivity in team formation and recognize when existing collaboration networks may exclude relevant expertise
- Develop of tools for efficient distribution of new funding opportunities, especially those outside of typical practices
- Create a database of Berkeley Lab researchers that is searchable by keyword and discipline
- Identify strategic themes, potential researchers, and researchers available to participate in new projects (for example, assess ahead of time if certain researchers are already overburdened)
- Provide a platform for researchers to discuss potential ideas outside of videoconferencing and meetings (for example, Slack channels were useful for understanding the early development of concepts for NVBL research)
- Establish a data sharing framework and infrastructure that can be used for new collaborations
- Engage contracting and technology transfer staff at team formation for rapid turnaround on materials transfer agreements, intellectual property agreements, and non-disclosure agreements.

While the establishment of the NVBL may have been a rare occurrence and required tactical decision making, the suggestions from the visioning session participants could be considered for future large scientific challenges, including those that are part of the DOE mission for scientific research.

21-BAO-3040